GenCore version 6.3
Copyright (c) 1993 - 2010 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2010, 20:11:08 ; Search time 1293 Seconds
(without alignments)
93164.247 Million cell updates/sec

Title: US-10-593-202-1_COPY_54755_55810
Perfect score: 1056
Sequence: 1 cacttccagcttcctttatc.....catccctggagtctctgcag 1056

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21196085 seqs, 57036634027 residues

Total number of hits satisfying chosen parameters: 42392170

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_bct:*
2: gb_env:*
3: gb_htc:*
4: gb_htg1:*

4: gb_htg1:*
5: gb_htg2:*
6: gb_htg3:*
7: gb_inv:*
8: gb_mam:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_phg:*
12: gb_pln:*
13: gb_pri:*
14: gb_rod:*
15: gb_sts:*
16: gb_syn:*
17: gb_tsa:*
18: gb_una:*
19: gb_vrl:*
20: gb_vrt:*

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
	1	1056	100.0	11928	9	CS175724	CS175724 Sequence
	2	1056	100.0	11928	9	CS433133	CS433133 Sequence
	3	1056	100.0	100140	9	CS175723	CS175723 Sequence
	4	1056	100.0	100140	9	CS433132	CS433132 Sequence
c	5	1056	100.0	185146	14	AC159967	AC159967 Mus muscu
c	6	1056	100.0	230474	14	AC102564	AC102564 Mus muscu
	7	1042.4	98.7	11539	9	CS175727	CS175727 Sequence
	8	1042.4	98.7	11539	9	CS433136	CS433136 Sequence
	9	1042.4	98.7	15913	14	AF235496S4	AF235499 Mus muscu
	10	896.2	84.9	938	14	AF189741	AF189741 Mus muscu
	11	547.4	51.8	560	9	CS175725	CS175725 Sequence
	12	547.4	51.8	560	9	CS433134	CS433134 Sequence

	12	547.4	51.8	560	9	CS433134	CS433134 Sequence
	13	454.4	43.0	247782	4	AC102987	AC102987 Rattus no
c	14	201.2	19.1	179418	4	AC141013	AC141013 Rattus no
	15	87.8	8.3	262016	4	AC111472	AC111472 Rattus no
	16	86.2	8.2	79922	14	BX004789	BX004789 Mouse DNA
	17	86	8.1	249266	4	AC130160	AC130160 Rattus no
c	18	84.4	8.0	231150	4	AC114195	AC114195 Rattus no
	19	84	8.0	3942	14	AF184913	AF184913 Mus muscu
	20	84	8.0	4125	9	CS433157	CS433157 Sequence
	21	84	8.0	4125	14	AF184912	AF184912 Mus muscu
	22	83	7.9	102716	5	AC161353	AC161353 Mus muscu
c	23	83	7.9	191823	14	AC116412	AC116412 Mus muscu
	24	82.4	7.8	229099	4	AC094076	AC094076 Rattus no
c	25	82.4	7.8	322690	4	AC116278	AC116278 Rattus no
	26	81.2	7.7	157999	4	AC117088	AC117088 Rattus no
c	27	81.2	7.7	277191	4	AC109077	AC109077 Rattus no
	28	80.6	7.6	260674	4	AC102979	AC102979 Rattus no
c	29	80.4	7.6	214895	14	AL845313	AL845313 Mouse DNA
c	30	79.8	7.6	50655	6	CU041229_3	Continuation (4 of
c	31	79.8	7.6	181720	6	CU041255	CU041255 Mus muscu
	32	79.8	7.6	226786	14	AL731896	AL731896 Mouse DNA
	33	79.8	7.6	247582	4	AC115339	AC115339 Rattus no
	34	79.6	7.5	189516	4	AC128065	AC128065 Rattus no
c	35	79.6	7.5	231235	4	AC093965	AC093965 Rattus no
	36	79.6	7.5	249474	4	AC136421	AC136421 Rattus no
	37	79.2	7.5	230384	4	AC141214	AC141214 Rattus no
	38	79.2	7.5	271139	4	AC109680	AC109680 Rattus no
	39	79	7.5	133485	14	AC118213	AC118213 Mus muscu
	40	78.8	7.5	236697	4	AC127845	AC127845 Rattus no
c	41	78.8	7.5	250101	4	AC106644	AC106644 Rattus no
c	42	78.4	7.4	132761	14	AL772225	AL772225 Mouse DNA
	43	78.4	7.4	181467	14	AL844586	AL844586 Mouse DNA
c	44	78.4	7.4	209901	4	AC102707	AC102707 Mus muscu
	45	77.4	7.3	166641	14	AC120210	AC120210 Rattus no

45 77.4 7.3 166641 14 AC120310 AC120310 Rattus no

ALIGNMENTS

RESULT 1
CS175724
LOCUS CS175724 11928 bp DNA linear PAT 12-OCT-2005
DEFINITION Sequence 2 from Patent WO2005090559.
ACCESSION CS175724
VERSION CS175724.1 GI:77627140
KEYWORDS .
SOURCE Mus musculus (house mouse)
  ORGANISM Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
  Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1
  AUTHORS Rohrschneider,L.R.
  TITLE Methods and compositions involving s-ship promoter regions
  JOURNAL Patent: WO 2005090559-A 2 29-SEP-2005;
  Fred Hutchinson Cancer Research Center (US)
FEATURES Location/Qualifiers
  source 1. .11928
  /organism="Mus musculus"
  /mol_type="unassigned DNA"
  /db_xref="taxon:10090"
ORIGIN

  Query Match 100.0%; Score 1056; DB 9; Length 11928;
  Best Local Similarity 100.0%;
  Matches 1056; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


Qy 1 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 60
Db 5626 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 5685
Qy 61 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 120
Db 5686 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 5745
Qy 121 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 180
Db 5746 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 5805
Qy 181 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 240
Db 5806 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 5865
Qy 241 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 300
Db 5866 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 5925
Qy 301 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 360
Db 5926 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 5985
Qy 361 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 420
Db 5986 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 6045
Qy 421 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 480
Db 6046 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 6105

```
Qy        481 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6106 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 6165

Qy        541 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACACCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6166 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACACCACA 6225

Qy        601 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6226 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 6285

Qy        661 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6286 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 6345

Qy        721 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6346 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 6405

Qy        781 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6406 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 6465

Qy        841 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6466 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 6525

Qy        901 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6526 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 6585

Qy        961 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 1020
              ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6586 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 6645

Qy       1021 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
                ||||||||||||||||||||||||||||||||||||
Db       6646 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 6681


RESULT 2
CS433133
LOCUS       CS433133               11928 bp    DNA     linear   PAT 19-OCT-2006
DEFINITION  Sequence 2 from Patent WO2006102188.
ACCESSION   CS433133
VERSION     CS433133.1  GI:116286367
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Rohrschneider,L.R.
  TITLE     Methods and compositions involving the s-ship promoter
  JOURNAL   Patent: WO 2006102188-A 2 28-SEP-2006;
            Fred Hutchinson Cancer Research Center (US)
FEATURES             Location/Qualifiers
     source          1. .11928
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

  Query Match             100.0%;  Score 1056;  DB 9;  Length 11928;
  Best Local Similarity   100.0%;
```

```
  Best Local Similarity   100.0%;
  Matches 1056;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5626 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 5685

Qy         61 CACCACACGGCCAATACTCCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5686 CACCACACGGCCAATACTCCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 5745

Qy        121 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5746 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 5805

Qy        181 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5806 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 5865

Qy        241 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5866 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 5925

Qy        301 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5926 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 5985

Qy        361 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5986 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 6045

Qy        421 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6046 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 6105
```

```
Db       6046 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 6105

Qy        481 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6106 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 6165

Qy        541 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACACCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6166 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACACCACA 6225

Qy        601 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6226 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 6285

Qy        661 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6286 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 6345

Qy        721 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6346 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 6405

Qy        781 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6406 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 6465

Qy        841 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6466 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 6525

Qy        901 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6526 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 6585
```

```
Qy         961 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6586 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 6645

Qy        1021 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
               |||||||||||||||||||||||||||||||||||||
Db        6646 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 6681


RESULT 3
CS175723
LOCUS       CS175723               100140 bp    DNA     linear   PAT 12-OCT-2005
DEFINITION  Sequence 1 from Patent WO2005090559.
ACCESSION   CS175723
VERSION     CS175723.1  GI:77627139
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Rohrschneider,L.R.
  TITLE     Methods and compositions involving s-ship promoter regions
  JOURNAL   Patent: WO 2005090559-A 1 29-SEP-2005;
            Fred Hutchinson Cancer Research Center (US)
FEATURES             Location/Qualifiers
     source          1. .100140
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN
```

```
  Query Match             100.0%;  Score 1056;  DB 9;  Length 100140;
  Best Local Similarity   100.0%;
  Matches 1056;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54755 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 54814

Qy         61 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54815 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 54874

Qy        121 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54875 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 54934

Qy        181 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54935 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 54994

Qy        241 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54995 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 55054

Qy        301 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55055 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 55114

Qy        361 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55115 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 55174

Qy        421 TGATGTGCGCTGCAGAGATGCTGAAGTGCGCAGCTGTGCGTTGATGAAGAAACGCTAG 480
```

```
Qy        421 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55175 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 55234

Qy        481 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55235 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 55294

Qy        541 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55295 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 55354

Qy        601 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55355 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 55414

Qy        661 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55415 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 55474

Qy        721 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55475 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 55534

Qy        781 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55535 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 55594

Qy        841 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55595 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 55654

Qy        901 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 960
              ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55655 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 55714

Qy        961 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55715 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 55774

Qy       1021 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
              |||||||||||||||||||||||||||||||||||
Db      55775 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 55810


RESULT 4
CS433132
LOCUS       CS433132               100140 bp    DNA     linear   PAT 19-OCT-2006
DEFINITION  Sequence 1 from Patent WO2006102188.
ACCESSION   CS433132
VERSION     CS433132.1  GI:116286366
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Rohrschneider,L.R.
  TITLE     Methods and compositions involving the s-ship promoter
  JOURNAL   Patent: WO 2006102188-A 1 28-SEP-2006;
            Fred Hutchinson Cancer Research Center (US)
FEATURES             Location/Qualifiers
     source          1. .100140
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
```

```
                     /db_xref="taxon:10090"
     misc_feature    27350. .78168
                     /note="N = A, C, G OR T/U"
ORIGIN

  Query Match             100.0%;  Score 1056;  DB 9;  Length 100140;
  Best Local Similarity   100.0%;
  Matches 1056;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54755 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 54814

Qy         61 CACCACACGGCCAATACTCCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54815 CACCACACGGCCAATACTCCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 54874

Qy        121 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54875 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 54934

Qy        181 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54935 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 54994

Qy        241 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54995 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 55054

Qy        301 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55055 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 55114

Qy        361 TAGTAGCATGCTTGCTCCAAGCAAACAGAGCAATGCCTCCACACTTACAATCTCACAAC 420
```

```
Qy        361 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55115 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 55174

Qy        421 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55175 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 55234

Qy        481 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55235 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 55294

Qy        541 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACACCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55295 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACACCACA 55354

Qy        601 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55355 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 55414

Qy        661 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55415 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 55474

Qy        721 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55475 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 55534

Qy        781 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55535 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 55594

Qy        841 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 900
              ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55595 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 55654

Qy        901 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55655 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 55714

Qy        961 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55715 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 55774

Qy       1021 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
                |||||||||||||||||||||||||||||||||||
Db      55775 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 55810


RESULT 5
AC159967/c
LOCUS       AC159967               185146 bp    DNA     linear   ROD 10-AUG-2005
DEFINITION  Mus musculus chromosome 1, clone RP23-302P9, complete sequence.
ACCESSION   AC159967
VERSION     AC159967.2  GI:72096116
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 185146)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 1, clone RP23-302P9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 185146)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.
```

AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (14-APR-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 185146)

AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,

Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (10-AUG-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA

COMMENT On Aug 10, 2005 this sequence version replaced gi:62543423.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
-------------- Project Information
Center project name: J5628
Center clone name: 302_P_9

Some of the sequence contained within base pairs 1 - 61047 was
stolen from accession AC102630
Some of the sequence contained within base pairs 75127 to the end
of the clone was stolen from accession AC102564

```
            of the clone was stolen from accession AC102564.
FEATURES             Location/Qualifiers
     source          1. .185146
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /map="1"
                     /clone="RP23-302P9"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     repeat_region   1. .72
                     /rpt_family="B1_Mur2"
     repeat_region   1088. .1438
                     /rpt_family="MTB"
     repeat_region   1477. .1583
                     /rpt_family="B1_Mus1"
     repeat_region   1937. .1984
                     /rpt_family="(TTTTTA)n"
     repeat_region   complement(1986. .2132)
                     /rpt_family="B1_Mus2"
     repeat_region   complement(2592. .2723)
                     /rpt_family="B1_Mur1"
     repeat_region   2730. .2753
                     /rpt_family="(T)n"
     repeat_region   complement(2755. .2895)
                     /rpt_family="B1_Mus2"
     repeat_region   complement(3848. .3974)
                     /rpt_family="B1_Mur3"
     repeat_region   complement(4415. .4490)
                     /rpt_family="LTR48"
     repeat_region   complement(6248. .6388)
                     /rpt_family="MER5B"
     repeat_region   7656. .7684
                     /rpt_family="(T)n"
```

```
                     /rpt_family="(T)n"
repeat_region        7744. .7799
                     /rpt_family="MIR"
repeat_region        8260. .8288
                     /rpt_family="(CA)n"
repeat_region        complement(8323. .9649)
                     /rpt_family="L1_Mus3"
repeat_region        10409. .10519
                     /rpt_family="MTD"
repeat_region        10518. .10617
                     /rpt_family="B1_Mur2"
repeat_region        11110. .11504
                     /rpt_family="MTB_Mm"
repeat_region        11625. .11789
                     /rpt_family="RSINE1"
repeat_region        11911. .12075
                     /rpt_family="Lx3B"
repeat_region        complement(12727. .12866)
                     /rpt_family="B1_Mur2"
repeat_region        complement(13372. .13559)
                     /rpt_family="RMER15"
repeat_region        14231. .14265
                     /rpt_family="(A)n"
repeat_region        complement(14284. .14404)
                     /rpt_family="B1F2"
repeat_region        complement(14806. .14900)
                     /rpt_family="PB1D7"
repeat_region        complement(15044. .15250)
                     /rpt_family="B3"
repeat_region        15689. .16181
                     /rpt_family="MT2_Mm"
repeat_region        16182. .16890
                     /rpt_family="MERVL-_Mm"
repeat_region        16886. .21308
```

```
repeat_region   16886. .21308
                /rpt_family="MERVL-_Mm"
repeat_region   21316. .21402
                /rpt_family="ORR1A3-int"
repeat_region   21403. .21440
                /rpt_family="MERVL-_Mm"
repeat_region   21441. .21933
                /rpt_family="MT2_Mm"
repeat_region   23018. .23181
                /rpt_family="B4A"
repeat_region   23247. .23413
                /rpt_family="MTEa"
repeat_region   23712. .23848
                /rpt_family="B1_Mur3"
repeat_region   23916. .24524
                /rpt_family="RMER1C"
repeat_region   complement(24525. .24836)
                /rpt_family="ORR1A4"
repeat_region   24837. .24886
                /rpt_family="RMER1C"
repeat_region   complement(24899. .25068)
                /rpt_family="Lx7"
repeat_region   25075. .25108
                /rpt_family="(CA)n"
repeat_region   complement(25115. .25612)
                /rpt_family="Lx7"
repeat_region   26642. .26722
                /rpt_family="RSINE1"
repeat_region   26808. .26828
                /rpt_family="GC_rich"
repeat_region   26812. .26832
                /rpt_family="GC_rich"
repeat_region   complement(27289. .27365)
                /rpt_family="MLT1L"
```

```
                     /rpt_family="MLT1L"
     repeat_region   27479. .27563
                     /rpt_family="(GA)n"
     repeat_region   27618. .27771
                     /rpt_family="RSINE1"
     repeat_region   27842. .28036
                     /rpt_family="ID_B1"
     repeat_region   28054. .28108
                     /rpt_family="ID4"
     repeat_region   complement(28312. .28635)
                     /rpt_family="MTE2b"
     repeat_region   complement(28626. .28706)
                     /rpt_family="MTE2b"
     repeat_region   28714. .28892
                     /rpt_family="B3"
     repeat_region   30190. .30456
                     /rpt_family="B4A"
     repeat_region   30479. .30514
                     /rpt_family="(CA)n"
     repeat_region   30898. .30933

  Query Match             100.0%;  Score 1056;  DB 14;  Length 185146;
  Best Local Similarity   100.0%;
  Matches 1056;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88181 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 88122

Qy         61 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88121 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 88062

Qy        121 TTATGTTGTATTGTTTATGATTATTATAGATATGTGTATATATGTGTGTGTATATATA 180
```

```
Qy        121 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88061 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 88002

Qy        181 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88001 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 87942

Qy        241 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87941 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 87882

Qy        301 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87881 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 87822

Qy        361 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87821 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 87762

Qy        421 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87761 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 87702

Qy        481 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87701 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 87642

Qy        541 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACACCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87641 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACACCACA 87582

Qy        601 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 660
              ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87581 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 87522

Qy        661 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87521 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 87462

Qy        721 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87461 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 87402

Qy        781 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87401 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 87342

Qy        841 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87341 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 87282

Qy        901 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87281 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 87222

Qy        961 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87221 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 87162

Qy       1021 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
                ||||||||||||||||||||||||||||||||||||
Db      87161 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 87126


RESULT 6
```

```
AC102564/c
LOCUS       AC102564               230474 bp    DNA     linear   ROD 21-JAN-2005
DEFINITION  Mus musculus chromosome 1, clone RP23-210C12, complete sequence.
ACCESSION   AC102564
VERSION     AC102564.7  GI:58000610
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 230474)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 1, clone RP23-210C12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 230474)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
```

Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 230474)

AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (10-DEC-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 4 (bases 1 to 230474)

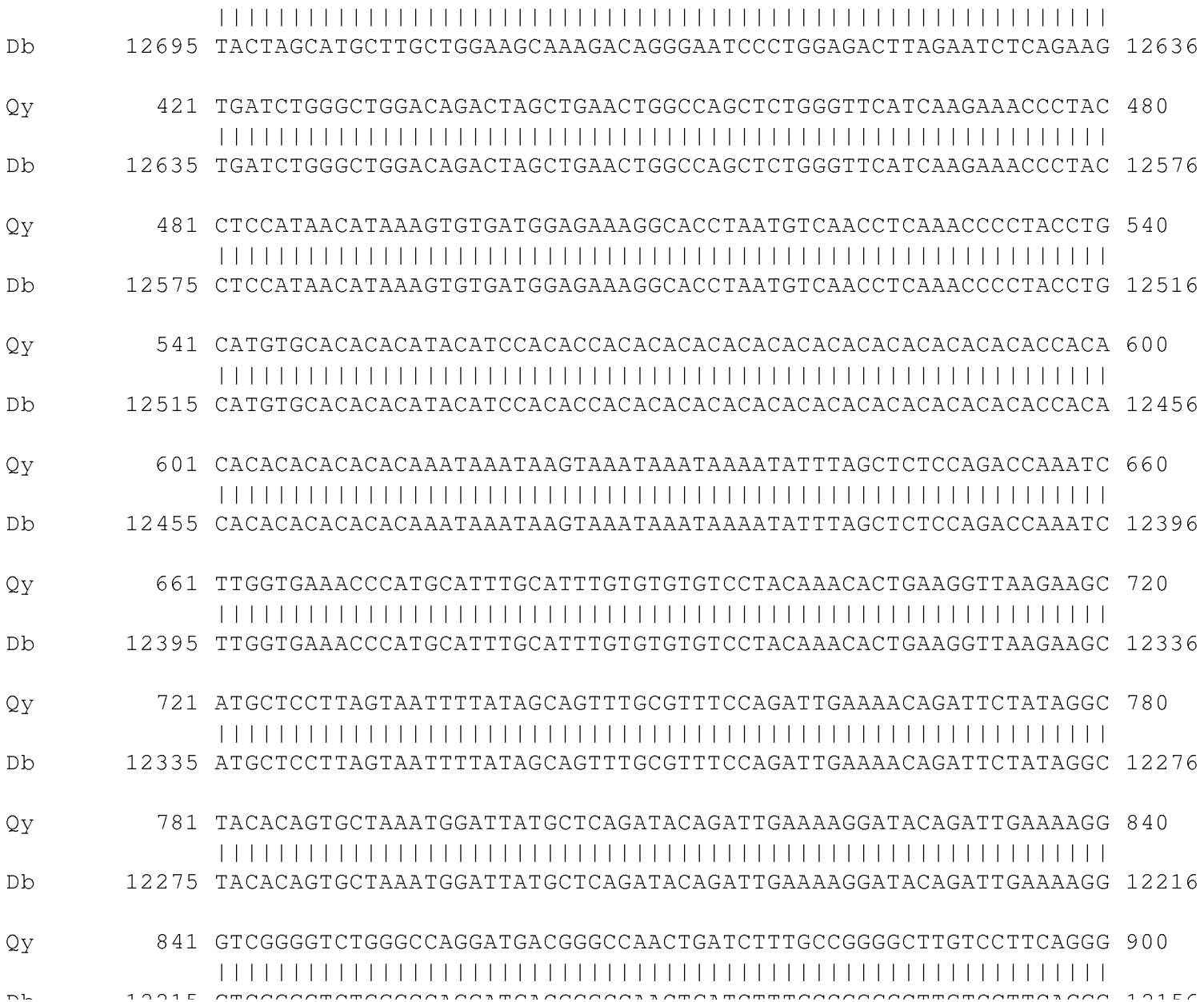
REFERENCE 4 (bases 1 to 230474)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jan 21, 2005 this sequence version replaced gi:55831552.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence-submissions@broad.mit.edu

```
                Contact: sequence_submissions@broad.mit.edu
              -------------- Project Information
                Center project name: L19015
                Center clone name: 210_C_12
              --------------.
FEATURES             Location/Qualifiers
     source          1. .230474
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /map="1"
                     /clone="RP23-210C12"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     repeat_region   complement(4. .80)
                     /rpt_family="PB1D9"
     repeat_region   complement(101. .307)
                     /rpt_family="B3"
     repeat_region   484. .519
                     /rpt_family="(CATA)n"
     repeat_region   657. .888
                     /rpt_family="B3A"
     repeat_region   960. .1002
                     /rpt_family="(TCCC)n"
     repeat_region   complement(1005. .1147)
                     /rpt_family="B1_MM"
     repeat_region   1306. .1387
                     /rpt_family="ID_B1"
     repeat_region   1468. .1491
                     /rpt_family="AT_rich"
     repeat_region   complement(2746. .2827)
                     /rpt_family="MIR"
     repeat_region   3070. .3313
                     /rpt_family="L2"
```

```
                     /rpt_family="L2"
repeat_region        complement(3536. .3591)
                     /rpt_family="MIR"
repeat_region        3985. .4358
                     /rpt_family="ORR1D"
repeat_region        5146. .5265
                     /rpt_family="RSINE1"
repeat_region        5266. .5297
                     /rpt_family="(CTCA)n"
repeat_region        5308. .5353
                     /rpt_family="(CAGA)n"
repeat_region        5537. .5727
                     /rpt_family="B2_Mm1"
repeat_region        5731. .5877
                     /rpt_family="B1_MM"
repeat_region        5880. .5949
                     /rpt_family="GA-rich"
repeat_region        6044. .6181
                     /rpt_family="B1_MM"
repeat_region        6182. .6218
                     /rpt_family="(CAAAA)n"
repeat_region        7488. .7586
                     /rpt_family="L2"
repeat_region        complement(7628. .7722)
                     /rpt_family="MIR"
repeat_region        8385. .8541
                     /rpt_family="B3A"
repeat_region        8548. .8590
                     /rpt_family="(TCTG)n"
repeat_region        8604. .8673
                     /rpt_family="B1F"
repeat_region        8674. .8731
                     /rpt_family="(GA)n"
repeat_region        8732. .8762
```

```
repeat_region   8732. .8762
                /rpt_family="ID_B1"
repeat_region   9094. .9219
                /rpt_family="MER34C"
repeat_region   complement(9374. .9492)
                /rpt_family="PB1"
repeat_region   complement(10096. .10217)
                /rpt_family="Lx6"
repeat_region   10498. .10747
                /rpt_family="B4A"
repeat_region   10748. .10870
                /rpt_family="RSINE1"
repeat_region   10897. .10923
                /rpt_family="(CAAA)n"
repeat_region   10979. .11028
                /rpt_family="(CATA)n"
repeat_region   11079. .11154
                /rpt_family="B1F"
repeat_region   11155. .11174
                /rpt_family="ID_B1"
repeat_region   11241. .11311
                /rpt_family="(TTTC)n"
repeat_region   complement(11313. .11471)
                /rpt_family="B1_MM"
repeat_region   12420. .12442
                /rpt_family="(TTTA)n"
repeat_region   12442. .12509
                /rpt_family="(TG)n"
repeat_region   complement(12519. .12803)
                /rpt_family="B4A"
repeat_region   12863. .12911
                /rpt_family="(TA)n"
repeat_region   13046. .13099
                /rpt_family="B1_MM"
```

```
                     /rpt_family="B1_MM"
     repeat_region   13106. .13263
                     /rpt_family="GA-rich"

  Query Match             100.0%;  Score 1056;  DB 14;  Length 230474;
  Best Local Similarity   100.0%;
  Matches 1056;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13055 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 12996

Qy         61 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12995 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 12936

Qy        121 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12935 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 12876

Qy        181 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12875 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 12816

Qy        241 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12815 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 12756

Qy        301 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12755 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 12696

Qy        361 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 420
              ''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
```


Db 12695 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 12636
Qy 421 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 480
Db 12635 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 12576
Qy 481 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 540
Db 12575 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 12516
Qy 541 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 600
Db 12515 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 12456
Qy 601 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 660
Db 12455 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 12396
Qy 661 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 720
Db 12395 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 12336
Qy 721 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 780
Db 12335 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 12276
Qy 781 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 840
Db 12275 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 12216
Qy 841 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 900

```
Db      12215 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 12156

Qy        901 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12155 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 12096

Qy        961 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12095 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 12036

Qy       1021 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
              |||||||||||||||||||||||||||||||||||
Db      12035 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 12000


RESULT 7
CS175727
LOCUS       CS175727               11539 bp    DNA     linear   PAT 12-OCT-2005
DEFINITION  Sequence 5 from Patent WO2005090559.
ACCESSION   CS175727
VERSION     CS175727.1  GI:77627143
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Rohrschneider,L.R.
  TITLE     Methods and compositions involving s-ship promoter regions
  JOURNAL   Patent: WO 2005090559-A 5 29-SEP-2005;
            Fred Hutchinson Cancer Research Center (US)
FEATURES             Location/Qualifiers
     source          1..11539
```

```
     source          1. .11539
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

  Query Match             98.7%;  Score 1042.4;  DB 9;  Length 11539;
  Best Local Similarity   99.8%;
  Matches 1054;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5271 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 5330

Qy         61 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5331 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 5390

Qy        121 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5391 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 5450

Qy        181 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5451 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 5510

Qy        241 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5511 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 5570

Qy        301 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5571 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 5630
```

```
Qy        361 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5631 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 5690

Qy        421 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5691 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 5750

Qy        481 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5751 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 5810

Qy        541 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5811 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 5870

Qy        601 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5871 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 5930

Qy        661 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5931 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 5990

Qy        721 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5991 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 6050

Qy        781 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6051 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 6110

Qy        841 CTGCGCGTGTCGCGCAGCATGAGCGCGAAGTGATGTTTGCGCGCGTTGTGCTTCAGCG 900
```

```
Qy        841 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 900
              |||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||
Db       6111 GTCGGGGTCTGGGCCAGGATGACGGGCCAACT-ATCTTTGCCCGGGCTTGTCCTTCAGGG 6169

Qy        901 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6170 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 6229

Qy        961 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6230 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 6289

Qy       1021 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
              |||||||||||||||||||||||||||||||||||
Db       6290 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 6325
```

```
RESULT 8
CS433136
LOCUS       CS433136               11539 bp    DNA     linear   PAT 19-OCT-2006
DEFINITION  Sequence 5 from Patent WO2006102188.
ACCESSION   CS433136
VERSION     CS433136.1  GI:116286370
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Rohrschneider,L.R.
  TITLE     Methods and compositions involving the s-ship promoter
  JOURNAL   Patent: WO 2006102188-A 5 28-SEP-2006;
            Fred Hutchinson Cancer Research Center (US)
```

```
            Fred Hutchinson Cancer Research Center (US)
FEATURES             Location/Qualifiers
     source          1. .11539
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
     misc_feature    1847
                     /note="N = A, C, G OR T/U"
ORIGIN

  Query Match             98.7%;  Score 1042.4;  DB 9;  Length 11539;
  Best Local Similarity   99.8%;
  Matches 1054;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5271 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 5330

Qy         61 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5331 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 5390

Qy        121 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5391 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 5450

Qy        181 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5451 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 5510

Qy        241 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5511 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 5570
```

```
Qy        301 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5571 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 5630

Qy        361 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5631 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 5690

Qy        421 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5691 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 5750

Qy        481 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5751 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 5810

Qy        541 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5811 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 5870

Qy        601 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5871 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 5930

Qy        661 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5931 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 5990

Qy        721 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5991 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 6050

Qy        781 TAGAGAGTGGTAAATGGATTATGGTGAGATAGAGATTGAAAAGGATAGAGATTGAAAAGC 840
```

```
Qy        781 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6051 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 6110

Qy        841 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 900
              |||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||
Db       6111 GTCGGGGTCTGGGCCAGGATGACGGGCCAACT-ATCTTTGCCCGGGCTTGTCCTTCAGGG 6169

Qy        901 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6170 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 6229

Qy        961 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6230 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 6289

Qy       1021 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
              |||||||||||||||||||||||||||||||||||
Db       6290 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 6325


RESULT 9
AF235496S4
LOCUS       AF235496S4             15913 bp    DNA     linear   ROD 18-OCT-2000
DEFINITION  Mus musculus SH2-containing inositol 5-phosphatase (Ship) gene,
            exons 3 through 6.
ACCESSION   AF235499
VERSION     AF235499.1  GI:10863172
KEYWORDS    .
SEGMENT     4 of 7
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
```

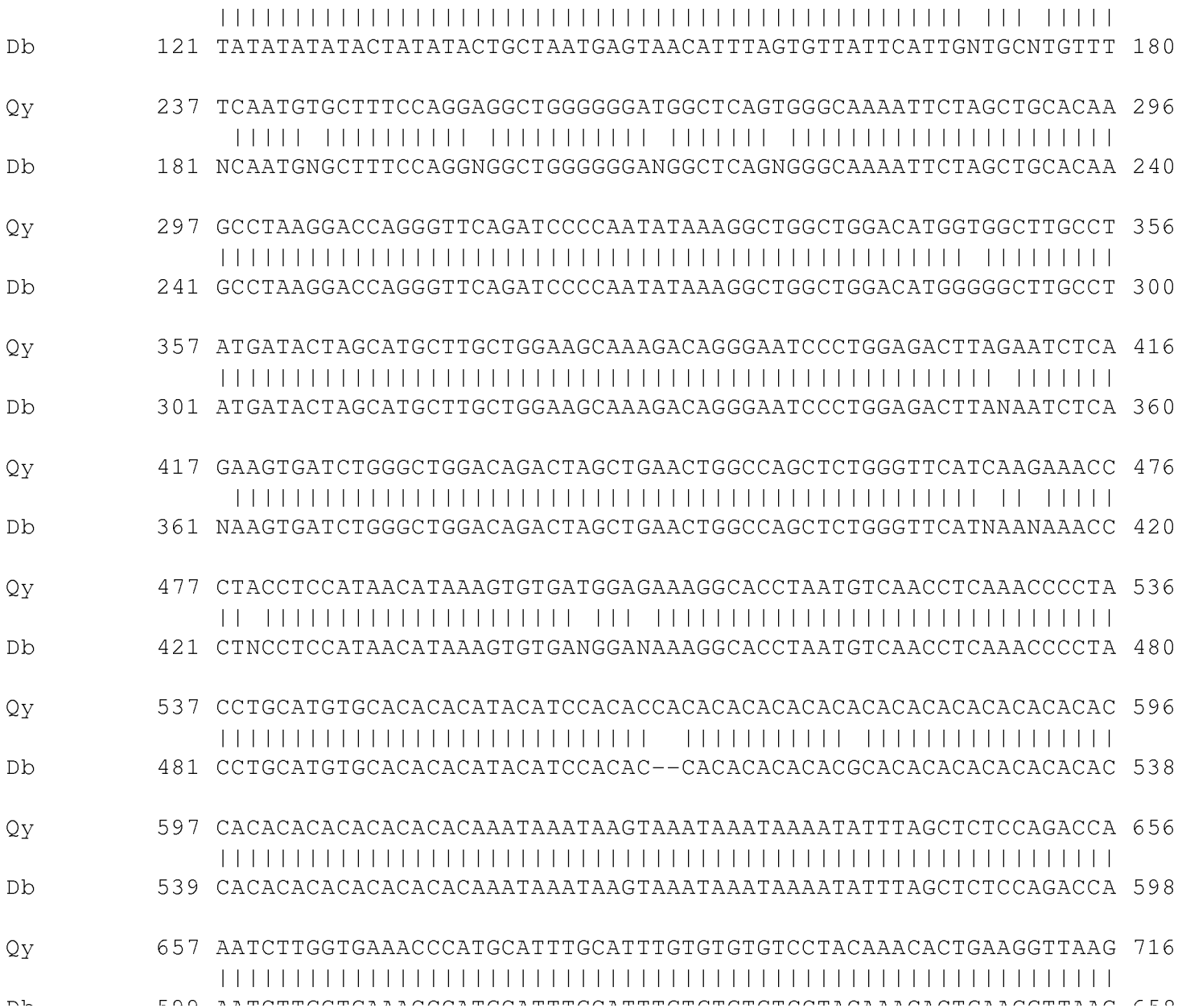
```
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 15913)
  AUTHORS   Wolf,I., Lucas,D.M., Algate,P.A. and Rohrschneider,L.R.
  TITLE     Cloning of the genomic locus of mouse SH2 containing inositol
            5-phosphatase (SHIP) and a novel 110-kDa splice isoform, SHIPdelta
  JOURNAL   Genomics 69 (1), 104-112 (2000)
   PUBMED   11013080
REFERENCE   2  (bases 1 to 15913)
  AUTHORS   Wolf,I.B., Lucas,D.M., Algate,P.A. and Rohrschneider,L.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2000) Division of Basic Sciences, Fred Hutchinson
            Cancer Research Center, 1100 Fairview Ave N, Seattle, WA
            98109-1024, USA
FEATURES             Location/Qualifiers
     source          1. .15913
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129/Sv"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /map="1C5"
     exon            3660. .3810
                     /gene="Ship"
                     /number=3
     exon            6166. .6337
                     /gene="Ship"
                     /number=4
     exon            7916. .8056
                     /gene="Ship"
                     /number=5
     exon            14555. .14642
                     /gene="Ship"
                     /number=6
```

```
                     /number=6
ORIGIN

  Query Match             98.7%;  Score 1042.4;  DB 14;  Length 15913;
  Best Local Similarity   99.8%;
  Matches 1054;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13540 CACTTCCAGCTTCCTTTATCATTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAAT 13599

Qy         61 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13600 CACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATTC 13659

Qy        121 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13660 TTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATATA 13719

Qy        181 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13720 TATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTTTCAA 13779

Qy        241 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13780 TGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGCCT 13839

Qy        301 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13840 AAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGCCTATGA 13899

Qy        361 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13900 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 13959
```

```
Db      13900 TACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCTCAGAAG 13959

Qy        421 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13960 TGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAACCCTAC 14019

Qy        481 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14020 CTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCCTACCTG 14079

Qy        541 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14080 CATGTGCACACACATACATCCACACCACACACACACACACACACACACACACACCACA 14139

Qy        601 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14140 CACACACACACACAAATAAATAAGTAAATAAATAAAATATTTAGCTCTCCAGACCAAATC 14199

Qy        661 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14200 TTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAGAAGC 14259

Qy        721 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14260 ATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTATAGGC 14319

Qy        781 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14320 TACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAAAAGG 14379

Qy        841 GTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTCAGGG 900
              |||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||
Db      14380 GTCGGGGTCTGGGCCAGGATGACGGGCCAACT-ATCTTTGCCCGGGCTTGTCCTTCAGGG 14438
```

```
Qy        901 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14439 AAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATTGCCT 14498

Qy        961 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14499 GGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGCTCCCAGGGAA 14558

Qy       1021 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
              |||||||||||||||||||||||||||||||||||
Db      14559 GTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 14594


RESULT 10
AF189741
LOCUS       AF189741                 938 bp    DNA     linear   ROD 18-OCT-2001
DEFINITION  Mus musculus SH2-containing inositol phosphatase SHIP gene, partial
            sequence.
ACCESSION   AF189741
VERSION     AF189741.1  GI:16223960
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 938)
  AUTHORS   Tu,Z., Ninos,J., Wang,J.W., Lemos,M. and Kerr,W.G.
  TITLE     An embryonic SHIP isoform generated by transcriptional initiation
            at an internal site participates in a different signaling complex
            than its hematopoietic counterpart
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 938)
```

```
REFERENCE   2  (bases 1 to 938)
  AUTHORS   Tu,Z. and Kerr,W.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-1999) Cell and Molecular Biology, University of
            Pennsylvania, 1318 Blockley Hall, 418 Guardian Drive, Philadelphia,
            PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .938
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            1. .>938
                     /gene="SH2-containing inositol phosphatase SHIP"
     promoter        1. .913
                     /gene="SH2-containing inositol phosphatase SHIP"
     exon            914. .>938
                     /gene="SH2-containing inositol phosphatase SHIP"
                     /number=1
ORIGIN

  Query Match             84.9%;  Score 896.2;  DB 14;  Length 938;
  Best Local Similarity   97.9%;
  Matches  921;  Conservative    0;  Mismatches   17;  Indels    3;  Gaps    2;

Qy          57 AAATCACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAA 116
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 AAATCACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAA 60

Qy         117 ATTCTTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATA 176
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ATTCTTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATA 120

Qy         177 TATATATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCATTGTTGCATGTTT 236
               ||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||||
```



```
Db        599 AATCTTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAGGTTAAG 658

Qy        717 AAGCATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTAT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 AAGCATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGATTCTAT 718

Qy        777 AGGCTACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAA 836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 AGGCTACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAGATTGAA 778

Qy        837 AAGGGTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTGTCCTTC 896
              |||||||||||||||||||||||||||||||||| ||||||||| ||||||||||||
Db        779 AAGGGTCGGGGTCTGGGCCAGGATGACGGGCCAACT-ATCTTTGCCCGGGCTTGTCCTTC 837

Qy        897 AGGGAAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATT 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 AGGGAAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACCTGAATT 897

Qy        957 GCCTGGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTAC 997
              |||||||||||||||||||||||||||||||||||||||||
Db        898 GCCTGGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTAC 938


RESULT 11
CS175725
LOCUS       CS175725                 560 bp    DNA     linear   PAT 12-OCT-2005
DEFINITION  Sequence 3 from Patent WO2005090559.
ACCESSION   CS175725
VERSION     CS175725.1  GI:77627141
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
```

```
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Rohrschneider,L.R.
  TITLE     Methods and compositions involving s-ship promoter regions
  JOURNAL   Patent: WO 2005090559-A 3 29-SEP-2005;
            Fred Hutchinson Cancer Research Center (US)
FEATURES             Location/Qualifiers
     source          1. .560
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

  Query Match             51.8%;  Score 547.4;  DB 9;  Length 560;
  Best Local Similarity   99.6%;
  Matches  559;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy        456 CTCTGGGTTCATCAAGAAACCCTACCTCCATAACATAAAGTGTGATGGAGAAAGGCACCT 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCTGGGTTCATCAAGAAACCCTACCTCCATAACATAAAGTGTGATGGAGAAAGGCACCT 60

Qy        516 AATGTCAACCTCAAACCCCTACCTGCATGTGCACACACATACATCCACACCACACACACA 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATGTCAACCTCAAACCCCTACCTGCATGTGCACACACATACATCCACACCACACACACA 120

Qy        576 CACACACACACACACACACACCACACACACACACACACAAATAAATAAGTAAATAAATAA 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACACACACACACACACACACCACACACACACACACACAAATAAATAAGTAAATAAATAA 180

Qy        636 AATATTTAGCTCTCCAGACCAAATCTTGGTGAAACCCATGCATTTGCATTTGTGTGTGTC 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATATTTAGCTCTCCAGACCAAATCTTGGTGAAACCCATGCATTTGCATTTGTGTGTGTC 240
```

```
Qy        696 CTACAAACACTGAAGGTTAAGAAGCATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTC 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTACAAACACTGAAGGTTAAGAAGCATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTC 300

Qy        756 CAGATTGAAAACAGATTCTATAGGCTACACAGTGCTAAATGGATTATGCTCAGATACAGA 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGATTGAAAACAGATTCTATAGGCTACACAGTGCTAAATGGATTATGCTCAGATACAGA 360

Qy        816 TTGAAAAGGATACAGATTGAAAAGGGTCGGGGTCTGGGCCAGGATGACGGGCCAACTGAT 875
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        361 TTGAAAAGGATACAGATTGAAAAGGGTCGGGGTCTGGGCCAGGATGACGGGCCAACT-AT 419

Qy        876 CTTTGCCGGGGCTTGTCCTTCAGGGAAGGGTTACAGGATTCACCACTGGGGTGTGGCCTA 935
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CTTTGCCCGGGCTTGTCCTTCAGGGAAGGGTTACAGGATTCACCACTGGGGTGTGGCCTA 479

Qy        936 TCTGCTGTTAGGACCTGAATTGCCTGGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTT 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TCTGCTGTTAGGACCTGAATTGCCTGGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTT 539

Qy        996 ACCTTGAACCTCTGCTCCCAG 1016
              |||||||||||||||||||||
Db        540 ACCTTGAACCTCTGCTCCCAG 560
```

RESULT 12
CS433134
LOCUS   CS433134   560 bp   DNA   linear   PAT 19-OCT-2006
DEFINITION Sequence 3 from Patent WO2006102188.
ACCESSION   CS433134
VERSION   CS433134.1 GI:116286368
KEYWORDS   .
SOURCE   Mus musculus (house mouse)

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Rohrschneider,L.R.
  TITLE     Methods and compositions involving the s-ship promoter
  JOURNAL   Patent: WO 2006102188-A 3 28-SEP-2006;
            Fred Hutchinson Cancer Research Center (US)
FEATURES             Location/Qualifiers
     source          1. .560
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

  Query Match             51.8%;  Score 547.4;  DB 9;  Length 560;
  Best Local Similarity   99.6%;
  Matches  559;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy        456 CTCTGGGTTCATCAAGAAACCCTACCTCCATAACATAAAGTGTGATGGAGAAAGGCACCT 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCTGGGTTCATCAAGAAACCCTACCTCCATAACATAAAGTGTGATGGAGAAAGGCACCT 60

Qy        516 AATGTCAACCTCAAACCCCTACCTGCATGTGCACACACATACATCCACACCACACACACA 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATGTCAACCTCAAACCCCTACCTGCATGTGCACACACATACATCCACACCACACACACA 120

Qy        576 CACACACACACACACACACACCACACACACACACACACAAATAAATAAGTAAATAAATAA 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACACACACACACACACACACCACACACACACACACACAAATAAATAAGTAAATAAATAA 180

Qy        636 AATATTTACGTGTCGACGCAAATGTTCGTCAAACGCATCGATTTCGATTTCGTCGTCGTC 695
```

```
Qy        636 AATATTTAGCTCTCCAGACCAAATCTTGGTGAAACCCATGCATTTGCATTTGTGTGTGTC 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATATTTAGCTCTCCAGACCAAATCTTGGTGAAACCCATGCATTTGCATTTGTGTGTGTC 240

Qy        696 CTACAAACACTGAAGGTTAAGAAGCATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTC 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTACAAACACTGAAGGTTAAGAAGCATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTC 300

Qy        756 CAGATTGAAAACAGATTCTATAGGCTACACAGTGCTAAATGGATTATGCTCAGATACAGA 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGATTGAAAACAGATTCTATAGGCTACACAGTGCTAAATGGATTATGCTCAGATACAGA 360

Qy        816 TTGAAAAGGATACAGATTGAAAAGGGTCGGGGTCTGGGCCAGGATGACGGGCCAACTGAT 875
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        361 TTGAAAAGGATACAGATTGAAAAGGGTCGGGGTCTGGGCCAGGATGACGGGCCAACT-AT 419

Qy        876 CTTTGCCGGGGCTTGTCCTTCAGGGAAGGGTTACAGGATTCACCACTGGGGTGTGGCCTA 935
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CTTTGCCCGGGCTTGTCCTTCAGGGAAGGGTTACAGGATTCACCACTGGGGTGTGGCCTA 479

Qy        936 TCTGCTGTTAGGACCTGAATTGCCTGGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTT 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TCTGCTGTTAGGACCTGAATTGCCTGGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTT 539

Qy        996 ACCTTGAACCTCTGCTCCCAG 1016
              |||||||||||||||||||||
Db        540 ACCTTGAACCTCTGCTCCCAG 560
```

RESULT 13
AC102987
LOCUS   AC102987   247782 bp   DNA   linear   HTG 10-MAY-2003
DEFINITION Rattus norvegicus clone CH230-52N2, WORKING DRAFT SEQUENCE, 3
            unordered pieces.

```
            unordered pieces.
ACCESSION   AC102987
VERSION     AC102987.5  GI:30521881
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 247782)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
```

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 247782)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (26-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

```
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 247782)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:23264611.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GHEE
                Center clone name: CH230-52N2
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 237739 bases at least Q40
                Consensus quality: 240323 bases at least Q30
                Consensus quality: 242042 bases at least Q20
                Estimated insert size: 247461; sum of contigs estimation
```

```
                Estimated insert size: 247461; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   244348: contig of 244348 bp in length
            *   244349   244448: gap of unknown length
            *   244449   246420: contig of 1972 bp in length
            *   246421   246520: gap of unknown length
            *   246521   247782: contig of 1262 bp in length.
FEATURES             Location/Qualifiers
     source          1. .247782
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-52N2"
     misc_feature    1. .1001
                     /note="wgs_end_extension;
                     clone_end:T7"
     misc_feature    2793. .3407
                     /note="clone_boundary;
                     clone_end:T7;
                     site:EcoRI;
                     end_sequence:BH318096"
     misc_feature    complement(187978. .243289)
                     /note="clone_boundary;
```

```
                     /note="clone_boundary;
                     clone_end:Sp6;
                     site:EcoRI;
                     end_sequence:BH318098"
     gap             244349. .244448
                     /estimated_length=unknown
     gap             246421. .246520
                     /estimated_length=unknown
ORIGIN

  Query Match             43.0%;  Score 454.4;  DB 4;  Length 247782;
  Best Local Similarity   72.0%;
  Matches  768;  Conservative    0;  Mismatches  236;  Indels   62;  Gaps   11;

Qy          1 CACTTCCAGCTTCCTTTATCA-TTTTAAAAAGAATTTCCTATGTGACTACTGTATTTAAA 59
              || ||||||||||||||| | |||||||||||| || ||||||   ||||||||||||
Db     146114 CAATTCCAGCTTCCTTTATTATTTTTAAAAAGAATATCATATGTGGGCACTGTATTTAAA 146173

Qy         60 TCACCACACGGCCAATACTCCCCCCCAACTCCTCCCAAATCCCCTCTACCCACTCAAATT 119
              |||   | |  ||  |||| ||| |||||||||||||||||||||||| ||||||||||
Db     146174 TCATTTCCCCTCCTGTACT-CCCTCCAACTCCTCCCAAATCCCCTCTACTCACTCAAATT 146232

Qy        120 CTTATCTTGTATTCTTTATCATTATTATACATATGTGTATATATGTGTGTGTGTATATAT 179
                | |  |||||| ||||| ||| ||| |  |||||  |||||  | | |
Db     146233 TCTGTTGTGTATTATTTATTATTGTTACATGTATGTACATATAAATATAT---------- 146282

Qy        180 ATATATACTATATACTGCTAATGAGTAACATTTAGTGTTATTCAT-TGTTGCATGTTTTC 238
                    |   ||  |||| |||||||||||||||||||| |||| ||||||||||||
Db     146283 -----AAACGTAATCTGCCAATGAGTAACATTTAGTGTTACTCATATGTTGCATGTTTTA 146337

Qy        239 AATGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCACAAGC 298
              |||||||||||||||||||||| |||||||||||||| ||| |||| ||||||||| |
Db     146338 AATGTGCTTTCCAGGAGGCTGGGGAGATGGCTCAGTGGGTAAAGTTCTTGCTGCACAAAC 146397
```


Qy 299 CTAAGGACCAGGGTTCAGATCCCCA------ATATAAAGGCTGGCTGGACATGGTGGCTT 352
Db 146398 CTAAGGACCAGGGTTCAGATCCCCAATCAGCATAGAAAAGCTGGCTGGATGTGGTGGCTC 146457
Qy 353 GCCTATGATACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAAT 412
Db 146458 TCCTATGATACTAGCGTGCTTGTTGGGAGCAAAGACGGGGAATCCCTAAAGACATGGAAT 146517
Qy 413 CTCAGAAGTGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGA 472
Db 146518 CTCTGAAGTGAGCT-GGCTGGACAGACTAGCTGAATCGACCAGCTCTGGGTTCATCAAGA 146576
Qy 473 AACCCTACCTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACC 532
Db 146577 AACCCTACCTCCATCA----------AGAGCGTGATCACACCTAATGACAGCCTCAGGCC 146626
Qy 533 CCTACCTGCATGTGCACACACATACATCCACACCACACACACACACACACACACACAC 592
Db 146627 CCTACATGCATGTGCACACACATACATGCACACACACACAAATAAGTTAATGAATAAAAT 146686
Qy 593 ACACCACACACACACACACACAAATAAATAAGTAAAT--AAATAAAATATTTAGCTCTCC 650
Db 146687 ATTTGGCTCTCCAGACCATAGTGAAACCTGGGCATTTGCATTTTTAGTGTGTGTGTGTGT 146746
Qy 651 AGACCAAATCTTGGTGAAACCCATGCATTTGCATTTGTGTGTGTCCTACAAACACTGAAG 710
Db 146747 GTGTGTGTGTGTGTGTGTGTGCATGTGTGTGTGTGTGTGTGTGTGTGTGTCTGAAA 146806
Qy 711 GTTAAGAAGCATGCTCCTTAGTAATTTTATAGCAGTTTGCGTTTCCAGATTGAAAACAGA 770
Db 146807 TTTAAGAAATATGCTCCTAAGTAATTTTATAGTTGTTTGTGTTTCCAGATTAGAAACAGA 146866

```
Qy        771 TTCTATAGGCTACACAGTGCTAAATGGATTATGCTCAGATACAGATTGAAAAGGATACAG 830
              ||||||||| |||| ||||    |||   |                        | |||||
Db     146867 TTCTATAGGTTACATAGTGGATTATGCTCT----------------------GGTACAG 146903

Qy        831 ATTGAAAAGGGTCGGGGTCTGGGCCAGGATGACGGGCCAACTGATCTTTGCCGGGGCTTG 890
              |||||| |||||| |||||| ||||||  ||| ||||   |||||||||||| |||| ||
Db     146904 ATTGAAGAGGGTCAGGGTCTAGGCCAGCCTGATGGGCTGGCTGATCTTTGCCCGGGCATG 146963

Qy        891 TCCTTCAGGGAAGGGTTACAGGATTCACCACTGGGGTGTGGCCTATCTGCTGTTAGGACC 950
              ||||| ||||||||| ||  ||||| || |  ||||| ||||| ||||| |||||| ||
Db     146964 TCCTTAAGGGAAGGGCTAAGGGATTGGCCGC-CGGGTGGGGCCTGTCTGCAGTTAGGGCC 147022

Qy        951 TGAATTGCCTGGAGTGTTTCTAGTTCCCACTAGTTGTTGAACTTTACCTTGAACCTCTGC 1010
              || ||| |||   ||| | ||||||| ||||| ||||||||||||||||||  |||| ||
Db     147023 TGTATT-CCTCAGGTGCTACTAGTTCGCACTAATTGTTGAACTTTACCTTGCTCCTCGGC 147081

Qy       1011 TCCCAGGGAAGTCATCAGGACTCTGCCATCCCTGGAGTCTCTGCAG 1056
              || ||| |||||||||||||| || || ||||||||||||||||||
Db     147082 TCTCAGAGAAGTCATCAGGACCCTCCCGTCCCTGGAGTCTCTGCAG 147127


RESULT 14
AC141013/c
LOCUS       AC141013               179418 bp    DNA     linear   HTG 27-MAR-2003
DEFINITION  Rattus norvegicus clone CH230-195L1, *** SEQUENCING IN PROGRESS
            ***, 67 unordered pieces.
ACCESSION   AC141013
VERSION     AC141013.1  GI:28875872
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidae; Muridae; Murinae; Rattus
```

Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 179418)

AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., Mcleod,M., Mcneill,T., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G.,

Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 179418)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (07-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 179418)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (27-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT ---------- Genome Center

Center: Baylor College of Medicine

Center code: BCM

```
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
    Center project name: GXKG
    Center clone name: CH230-195L1
-------------- Summary Statistics
    Sequencing vector: Plasmid;
    Chemistry: Dye-terminator Big Dye: 100% of reads
    Assembly program: Phrap; version 0.990329
    Consensus quality: 139634 bases at least Q40
    Consensus quality: 148988 bases at least Q30
    Consensus quality: 156753 bases at least Q20
    Estimated insert size: 145919; sum-of-contigs estimation
    Quality coverage: 2x in Q20 bases; sum-of-contigs estimation
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 67 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1465: contig of 1465 bp in length
*     1466     1565: gap of unknown length
*     1566     2744: contig of 1179 bp in length
*     2745     2844: gap of unknown length
*     2845     3976: contig of 1132 bp in length
*     3977     4076: gap of unknown length
*     4077     5663: contig of 1587 bp in length
*     5664     5763: gap of unknown length
```

```
*      5664      5763: gap of unknown length
*      5764      7208: contig of 1445 bp in length
*      7209      7308: gap of unknown length
*      7309      8597: contig of 1289 bp in length
*      8598      8697: gap of unknown length
*      8698      9901: contig of 1204 bp in length
*      9902     10001: gap of unknown length
*     10002     11343: contig of 1342 bp in length
*     11344     11443: gap of unknown length
*     11444     12518: contig of 1075 bp in length
*     12519     12618: gap of unknown length
*     12619     14117: contig of 1499 bp in length
*     14118     14217: gap of unknown length
*     14218     15454: contig of 1237 bp in length
*     15455     15554: gap of unknown length
*     15555     17016: contig of 1462 bp in length
*     17017     17116: gap of unknown length
*     17117     18499: contig of 1383 bp in length
*     18500     18599: gap of unknown length
*     18600     19774: contig of 1175 bp in length
*     19775     19874: gap of unknown length
*     19875     21219: contig of 1345 bp in length
*     21220     21319: gap of unknown length
*     21320     22672: contig of 1353 bp in length
*     22673     22772: gap of unknown length
*     22773     24752: contig of 1980 bp in length
*     24753     24852: gap of unknown length
*     24853     26001: contig of 1149 bp in length
*     26002     26101: gap of unknown length
*     26102     28650: contig of 2549 bp in length
*     28651     28750: gap of unknown length
*     28751     30694: contig of 1944 bp in length
*     30695     30794: gap of unknown length
*     30795     32634: contig of 1840 bp in length
```

```
*     30795    32634: contig of 1840 bp in length
*     32635    32734: gap of unknown length
*     32735    35068: contig of 2334 bp in length
*     35069    35168: gap of unknown length
*     35169    36796: contig of 1628 bp in length
*     36797    36896: gap of unknown length
*     36897    38190: contig of 1294 bp in length
*     38191    38290: gap of unknown length
*     38291    39461: contig of 1171 bp in length
*     39462    39561: gap of unknown length
*     39562    41210: contig of 1649 bp in length
*     41211    41310: gap of unknown length
*     41311    43224: contig of 1914 bp in length
*     43225    43324: gap of unknown length
*     43325    45608: contig of 2284 bp in length
*     45609    45708: gap of unknown length
*     45709    47065: contig of 1357 bp in length
*     47066    47165: gap of unknown length
*     47166    48454: contig of 1289 bp in length
*     48455    48554: gap of unknown length
*     48555    50246: contig of 1692 bp in length
*     50247    50346: gap of unknown length
*     50347    52037: contig of 1691 bp in length
*     52038    52137: gap of unknown length
*     52138    53888: contig of 1751 bp in length
*     53889    53988: gap of unknown length
*     53989    55344: contig of 1356 bp in length
*     55345    55444: gap of unknown length
*     55445    56683: contig of 1239 bp in length
*     56684    56783: gap of unknown length
*     56784    58824: contig of 2041 bp in length
*     58825    58924: gap of unknown length
*     58925    60588: contig of 1664 bp in length
*     60589    60688: gap of unknown length
```

```
              *    60589    60688: gap of unknown length
              *    60689    63023: contig of 2335 bp in length
              *    63024    63123: gap of unknown length
              *    63124    65633: contig of 2510 bp in length
              *    65634    65733: gap of unknown length
              *    65734    67706: contig of 1973 bp in length
              *    67707    67806: gap of unknown length
              *    67807    69538: contig of 1732 bp in length
              *    69539    69638: gap of unknown length
              *    69639    71035: contig of 1397 bp in length
              *    71036    71135: gap of unknown length
              *    71136    72415: contig of 1280 bp in length
              *    72416    72515: gap of unknown length
              *    72516    75179: contig of 2664 bp in length
              *    75180    75279: gap of unknown length
              *    75280    77087: contig of 1808 bp in length
              *    77088    77187: gap of unknown length
              *    77188    78940: contig of 1753 bp in length
              *    78941    79040: gap of unknown length
              *    79041    81122: contig of 2082 bp in length
              *    81123    81222: gap of unknown length
              *    81223    83707: contig of 2485 bp in length
              *    83708    83807: gap of unknown length
              *    83808    86713: contig of 2906 bp in length
              *    86714    86813: gap of unknown length

  Query Match             19.1%;  Score 201.2;  DB 4;  Length 179418;
  Best Local Similarity   81.0%;
  Matches  277;  Conservative    0;  Mismatches   48;  Indels   17;  Gaps    3;

Qy        230 CATGTTTTCAATGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGC 289
              ||   |||   |||||||||||||||||||||| |||||||||||||| ||| |||| ||
Db      11317 CAATGTTTTTCTGTGCTTTCCAGGAGGCTGGGGAGATGGCTCAGTGGGTAAAGTTCTTGC 11258
```

```
Qy        290 TGCACAAGCCTAAGGACCAGGGTTCAGATCCCCA------ATATAAAGGCTGGCTGGACA 343
              ||||||| ||||||||||||||||||||||||      ||| ||| ||||||||||
Db      11257 TGCACAAACCTAAGGACCAGGGTTCAGATCCCCAATCAGCATAGAAAAGCTGGCTGGATG 11198

Qy        344 TGGTGGCTTGCCTATGATACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAG 403
              ||||||||  |||||||||||||| |||||| ||| ||||||||| ||||||||||  ||
Db      11197 TGGTGGCTCTCCTATGATACTAGCGTGCTTGTTGGGAGCAAAGACGGGGAATCCCTAAAG 11138

Qy        404 ACTTAGAATCTCAGAAGTGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGT 463
              || | ||||||| ||||||| || |||||||||||||| |||||  | ||||||||||||
Db      11137 ACATGGAATCTCTGAAGTGAGCT-GGCTGGACAGACTATCTGAATCGACCAGCTCTGGGT 11079

Qy        464 TCATCAAGAAACCCTACCTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAA 523
              ||||||||||||||||||||| |             |    |   ||||||||| ||
Db      11078 TCATCAAGAAACCCTACCTCCATCA----------AGAGCGTGATCACACCTAATGACAG 11029

Qy        524 CCTCAAACCCCTACCTGCATGTGCACACACATACATCCACAC 565
              |||||  ||||||| |||||||||||||||||||| |||||
Db      11028 CCTCACGCCCCTACATGCATGTGCACACACATACATGCACAC 10987
```

```
RESULT 15
AC111472
LOCUS       AC111472              262016 bp    DNA     linear   HTG 13-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-82D4, WORKING DRAFT SEQUENCE.
ACCESSION   AC111472
VERSION     AC111472.5  GI:30579364
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus
```

Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 262016)

AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

  TITLE Direct Submission

  JOURNAL Unpublished

REFERENCE 2 (bases 1 to 262016)

  AUTHORS Worley,K.C.

  TITLE Direct Submission

  JOURNAL Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 262016)

  AUTHORS Rat Genome Sequencing Consortium.

  TITLE Direct Submission

  JOURNAL Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On May 13, 2003 this sequence version replaced gi:24942244.

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
-------------- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
    Center project name: GMJZ
    Center clone name: CH230-82D4
-------------- Summary Statistics
    Assembly program: Atlas 3.0;
    Consensus quality: 245356 bases at least Q40
    Consensus quality: 247686 bases at least Q30
    Consensus quality: 249572 bases at least Q20
    Estimated insert size: 257162; sum-of-contigs estimation
    Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces

```
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1   262016: contig of 262016 bp in length.
FEATURES             Location/Qualifiers
     source          1. .262016
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-82D4"
     misc_feature    1. .1000
                     /note="wgs_contig"
     misc_feature    260556. .262016
                     /note="wgs_contig"
ORIGIN

  Query Match              8.3%;  Score 87.8;  DB 4;  Length 262016;
  Best Local Similarity   58.5%;
  Matches  210;  Conservative    0;  Mismatches  142;  Indels    7;  Gaps    3;

Qy        235 TTTCAATGTGCTTTCCAGGAGGCTGGGGGGATGGCTCAGTGGGCAAAATTCTAGCTGCAC 294
              | ||  ||    | ||||  ||   ||| ||||||  | |||| ||||  ||  ||  |
Db       4085 TCTCTTTGGATCTCCCAGAGGGGCTGGGTGATGGCATAATGGGTAAAAAGCTCTCTATAT 4144

Qy        295 AAGCCTAAGGACCAGGGTTCAGATCCCCAATATAAAGGCTGGCTGGACATGGTGGCTTGC 354
              ||||   |||||| | |||  ||||||||||  | | |        |  ||  |  ||
Db       4145 AAGCAGGAGGACCTGCGTTTGGATCCCCAATGCACATGTAAAAGCCAAGTGTAGTGTTCT 4204

Qy        355 CTATGATACTAGCATGCTTGCTGGAAGCAAAGACAGGGAATCCCTGGAGACTTAGAATCT 414
               ''      ' ' '   '''  '  ''  ''''   '''       ' '   ' ' '''
```

```
                ||      | | |   |||  |  ||  ||||   |||       | |   | | |||
Db       4205 GTAA-----TTGTAGCATTGAGGTCAGGGAAGAGGTGGAGCAGGATGGGGAGTGGGATCC 4259

Qy        415 CAGAAGTGATCTGGGCTGGACAGACTAGCTGAACTGGCCAGCTCTGGGTTCATCAAGAAA 474
              |           ||| | |  || |||  |||| |||  ||||||| ||||     ||| |
Db       4260 CTAGGTCTTGTAGGGATAGCTAGTCTA-TTGAAATGGGGAGCTCTAGGTTTGATGAGAGA 4318

Qy        475 CCCTACCTCCATAACATAAAGTGTGATGGAGAAAGGCACCTAATGTCAACCTCAAACCCC 534
              | || | || | ||   | ||  | || ||||||| |||| ||| || | |||  ||  |
Db       4319 CTCTGCTTCGAAAAGTAAGAGAATAATAGAGAAAGACACCAAATATCGATCTCTTACTTC 4378

Qy        535 TACCTGCATGTGCACACACATACATCCACACCACACACACACACACACACACACACA 593
               || ||     || |||||| |||  |||| ||||||||||||||||||||||||||
Db       4379 CACATGTGCACGCGCACACACACACACACA-CACACACACACACACACACACACACACA 4436
```

Search completed: June 16, 2010, 20:34:55
Job time : 1427 secs